DOCUMENT RESUME

ED 471 767 SE 067 026

AUTHOR	Johnson, Yolanda; Hofbauer, Pamela
TITLE	Describing Middle School Students' Organization of Statistical Data.
PUB DATE	2002-00-00
NOTE	11p.; In: Proceedings of the Annual Meeting [of the] North American Chapter of the International Group for the Psychology of Mathematics Education (24th, Athens, GA, October 26-29, 2002). Volumes 1-4; see SE 066 887.
AVAILABLE FROM	ERIC/CSMEE Publications, 1929 Kenny Road, Columbus, OH 43210-1080. Tel: 800-276-0462 (Toll Free).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Concept Formation; Data Processing; Junior High Schools; *Mathematics Education; Middle Schools; *Statistics

ABSTRACT

The purpose of this study was to describe how middle school students physically arrange and organize statistical data. A case-study analysis was used to define and characterize the styles in which students handle, organize, and group statistical data. A series of four statistical tasks (Mooney, Langrall, Hofbauer, & Johnson, 2001) were given to twelve students, 4 from each of grades 6-8, in an interview format. Based on an analysis of the interview data, five categories for students' arrangement strategies emerged: No Arrangement (students made no attempt to arrange the data); Clustered Arrangement (students sorted the given data in groups with no totals); Sequential Arrangement (students listed data in least to greatest order or alphabetical order); Summative Arrangement (students provided totals for categories or groups of data); and Regrouped-Summative Arrangement (students re-grouped the data into new categories and gave totals). This study implies that students need to be presented with raw, unorganized data sets in addition to organized data sets and ask students to arrange or rearrange data in a meaningful way. (Author)

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

D. Owens

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

DESCRIBING MIDDLE SCHOOL STUDENTS' ORGANIZATION OF STATISTICAL DATA

Yolanda Johnson
Illinois State University
yajohns@ilstu.edu

Pamela Hofbauer
Illinois State University
pshofba@ilstu.edu

The purpose of this study was to describe how middle school students physically arrange and organize statistical data. A case-study analysis was used to define and characterize the styles in which students handle, organize, and group statistical data. A series of four statistical tasks (Mooney,, Langrall, Hofbauer, & Johnson, 2001) were given to twelve students, 4 from each of grades 6–8, in an interview format. Based on an analysis of the interview data, five categories for students' arrangement strategies emerged: *No Arrangement* (students made no attempt to arrange the data); *Clustered Arrangement* (students sorted the given data in groups with no totals); *Sequential Arrangement* (students listed data in least to greatest order or alphabetical order); *Summative Arrangement* (students provided totals for categories or groups of data); and *Regrouped-Summative Arrangement* (students re-grouped the data into new categories and gave totals). This study implies that students need to be presented with raw, unorganized data sets in addition to organized data sets and ask students to arrange or rearrange data in a meaningful way.

ED 471 767

As part of the school curriculum, statistics has been a topic of research for several decades (Jones & Coxford, 1970; National Council of Teachers of Mathematics [NCTM], 2000; Pieters & Kinsella, 1959; Shulte & Smart, 1981). Paulos (1988) mentions that graphical representation is important for every day matters; while Steen (1997) points out that the centerpiece of middle school mathematics includes topics in exploratory data analysis and numerical reasoning. It is important for students to study statistics in the middle school, if not earlier, because the numbers that surround our daily lives are staggering. These numbers can be found on the Internet and in newspapers, television, movie ratings, food labels, consumer reports, and census reports to name a few. Making intellectual decisions in the global society in which we live is a task that every person needs to be able to perform accurately in order to be a better informed citizen and consumer.

Students come to understand the fundamentals of statistical ideas by being engaged in data handling processes such as collecting, organizing, and representing data (NCTM, 2000). According to NCTM, middle school students should begin to compare the effectiveness of various ways of organizing data for analysis or presentation (p. 49). It is also important to remember that in real-life situations, data are not always presented in an organized fashion; in fact, data can be messy. In order for students to make a connection between school and out-of-school mathematics, they should be encouraged to solve problem situations within a meaningful statistical con-

BEST COPY AVAILABLE

text. However, many of the current middle school mathematics textbooks do not give students the opportunity to handle raw data (e.g., Charles, et al., 2000). Students are typically given organized data within a real-world situation then told to construct a particular graph. While much of the statistics research at the middle school level has looked at how students deal with reducing, representing, and analyzing data (e.g., Curcio, 1987; Mokros & Russell, 1995; Strauss & Bichler, 1988), virtually no research has examined how students organize or reorganize data.

We have taken the view that the first step in building a knowledge base in this area involves describing or characterizing the general strategies that students use when working with data. Thus, this study was exploratory in nature and did not interpret the data through the lens of any particular theoretical perspective. The intent of this study was to describe the strategies middle school students use to organize data.

Method

Participants

Students in grades six through eight from four mid-western schools formed the population for this study. Twelve students, four from each grade level, were selected for case-study analysis based on levels of performance in mathematics: one high, two middle, and one low.

Instrument

The interview protocol (Mooney et al., 2001) was comprised of four tasks, each requiring students to arrange data to complete the task. For each task, the students were given the complete data set on paper and the same data with each data value on a separate card. Questions were designed so students could respond orally or by generating tables or graphical representations. At the end of each task, students were given the opportunity to describe an alternate method of reorganizing the data.

A description of each task is provided in Table 1. For each task, students were asked to arrange the data in a suitable format to be presented in the school newspaper. For Task 1, *Shoe Size,* the students were asked to arrange a set of 50 shoe sizes; for Task 2, *Teachers' Pets,* the students were given a set of 39 pets to organize. In Task 3, *Academy Awards*, the students were given the ages of each year's Best Actor and Best Actress Academy Award winners for 30 years and asked to rearrange the data to choose an appropriate headline. Finally, in Task 4, *Classical Music*, the students were given test scores for students in two math classes. The students were asked to reorganize the data to determine if the test-takers who listened to classical music while studying performed better than those who did not. It should be noted that the data for the first two tasks was unorganized. Data for the last two tasks were organized, but potentially needed to be reorganized to complete the tasks.

Table 1. Task Interview Protocol

Task 1: Shoe Sizes 50 eighth-grade students were surveyed about their shoe size. This list shows the data collected. Your job is to arrange the data to be presented in the school newspaper.
Task 2: Teachers' Pets The teachers at your school were asked what kinds of pets they have at home. In all, the teachers had 39 pets. A list of these pets is shown on this page. Your job is to arrange the data to be presented in the school newspaper.
Task 3: Oscar Winners This table shows the ages of the last 30 winners for the Best Actor and Best Actress in a movie. The editor of the school newspaper wants you to arrange the data to be presented in the school newspaper and to determine which of these 3 headlines should go with the data: Headline 1: Academy Likes Actors Older Than Actresses Headline 2: No Age Bias in Best Actor and Best Actress Winners Headline 3: Academy Likes Actresses Older Than Actors
Task 4: Study Habits Mrs. Jones talked to the students in her mathematics classes one day about an article she read. It said that children who listened to classical music while studying performed better on tests than children who did not listen to classical music while studying. Some of her students planned to listen to classical music while studying for the next math exam. The results of the 80-point test are listed on this table. The students who listened to classical music have an "X" marked next to their name. Your job is to arrange the data to see if students who listened to classical music while studying performed better on the math test than students who did not listen to classical music while studying. The editor of the school newspaper wants you to present the data along with a headline about the comparison

Procedure

Using the interview protocol, each student was individually interviewed during a 60-minute session. All interviews were audio taped and all student-generated work was collected. The interviews were transcribed for analysis.

Data Sources and Analysis

Data sources consisted of the transcribed interviews, students' written work and data displays, researcher field notes, and summaries generated during the analysis. Using a grounded theory approach, we examined students' responses to discern general patterns of arranging data. A double coding procedure (Miles & Huberman, 1994) was used to analyze student responses. We individually coded each student's response based on the type of arrangement the student produced. Then, we compared the coded responses within a category to discern the characteristics of that type of arrangement. Throughout this process differences in categorization and coding were discussed and agreement was negotiated.

Results

We found that students' arrangement strategies fit into five categories, which will be described in the following paragraphs. The resulting categories were as follows: No Arrangement, Clustered Arrangement, Sequential Arrangement, Summative Arrangement, or Regrouped-Summative Arrangement.

In *No Arrangement,* students made no attempt to arrange the data, leaving it as raw data, or they arranged it inappropriately (see Figure 1). The most obvious way to fit in this category was if no arrangement was provided. Two students (6B, 6C) in the Academy Awards task and two in the Classical Music task (6A, 6B) did not attempt to rearrange the data. The Classical Music task generated the most varied types of responses that we categorized as No Arrangement. Some students dealt with the unequal ends data in an inappropriate manner. Unequal ends are defined here as meaning the data did not include the same number of students in each group. As a result, they forced the data into data pairs, using only 13 of the non-classical music scores to match the 13 test-takers who listened to classical music (7A, 7D). Student 7A chose the top 13 scores of those who did not listen to classical music, whereas student 7D chose the first 13 scores he saw to compare to the test-takers who listened to classical music. The most common example of an inappropriate arrangement was when students found averages that either did not represent the data or did not solve the task. This occurred in the Shoe Size task with four students (6A, 7B, 8A, 8C), in the Academy Awards task with five students (7D, 8A, 8B, 8C, 8D) and in Classical Music task with five students (7B, 8A, 8B, 8C, 8D) who found the average shoe size, average ages of the actors/ actresses, and average test scores, respectively. Three other inappropriate arrangements were presented by students (6D, 7B, and 8C) solving the Classical Music task. Student 6D found the average of the top three scores and lowest

three scores of students who listened and did not listen to classical music and compared those averages. The test in the task was out of 80 possible points, so student 7B figured out what score was a 70% (which he figured, incorrectly, to be 57 out of 80) and found how many students scored above this score. One student (8C) stated that a graph would be inappropriate and that a description of the results would suffice in the school newspaper.

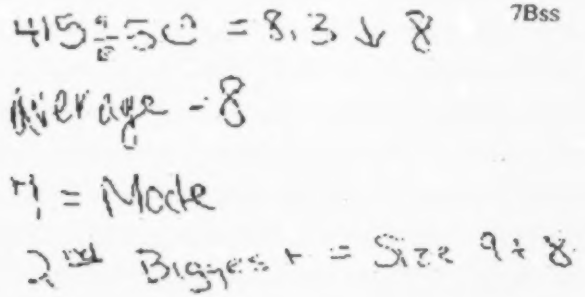

Figure 1. No Arrangement.

With *Clustered Arrangement*, students sorted the given data in groups with no totals (see Figure 2). This arrangement appeared only in the Teachers' Pets and Academy Awards tasks. For example, with Teachers' Pets, some students (6B, 6D, 7C, 8D) grouped the pets by the breed of animal or made a new group, but they did not provide a total amount. The students who provided graphical representations (6B and 6D) used a line plot with an 'X' representing each animal. In the Academy Awards task, student 6D placed a '+' beside actors and actresses who were 40 or over and a '−'beside those who were under 40. At the end she saw that she had more '+' for actors than actresses and chose Headline 1. One student (7A) subtracted the actor's age from the actress's each year to see if the difference was more than 10 years (Headline 3), less than 10 years (Headline 2), or negative (Headline 1). Still another student (8B) who provided a Clustered Arrangement looked at the years that the actresses were older to discover this occurred "in about one-third of the years".

In *Sequential Arrangement*, students listed data either in least to greatest or in alphabetical order (see Figure 3). For instance, with Shoe Size, six students (6A, 6B, 6D, 7C, 7D, 8C) decided to list the sizes in order from least to greatest. In the Teachers' Pets task, one student (6C) placed the animals in alphabetical order. For the Academy Awards task, four students (6A, 7C, 7D, and 8C) listed the data in numerical order. In addition, student 7C noticed that there were more shaded ages (actors) close to the

BEST COPY AVAILABLE

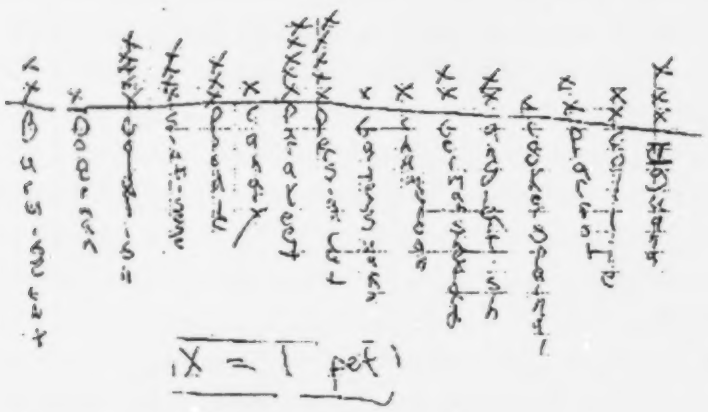

Figure 2. Clustered arrangement.

bottom, where the older ages were and less at the top, where the younger ages were, to reach the conclusion that Headline 1 would fit the data. Student 7D suggested to put the data in a graph with age on one axis and year on another but did not suggest the exact type of graph to represent this data. Four students (6C, 6D, 7C, 8C) arranged the cards in least to greatest order for the Classical Music task and expressed that arrangement of the cards was the best way to present the information in the school newspaper.

With *Summative Arrangement*, students provided totals for categories or groups of data (see Figure 4). For instance, with Teachers' Pets, a student sorted the pets by breed and then provided the total number of each breed in the data set. For the Shoe Size task, in particular, the students mostly arranged the data in a Summative Arrangement (6C, 7A, 7B, 8B, 8D), which included some variation of finding how many students wore certain shoe sizes and representing the data in a bar graph or line plot. One student (6C) placed the animals in alphabetical order, giving a Sequential Arrangement. When asked to provide an alternate arrangement, she joined four other students (7B, 7D, 8A, and 8B), who organized the data in a Summative Arrangement according to the breeds of animals (i.e., Doberman, Cocker Spaniel, German Shepherd, etc.). Two students (8A and 8D) used a Summative Arrangement, comparing age totals between the actors and actresses in the Academy Awards tasks.

Throughout the coding process, we found that some students regrouped the data into new categories and gave totals, which was characteristic of the *Regrouped-Summative Arrangement* (see Figure 5). For example, with Academy Awards, student 7B placed a mark next to winners who were 40 or older and then noted that 10 out of 30

BEST COPY AVAILABLE

actresses had marks whereas 19 out of 30 actors had marks leading to a Regrouped-Summative Arrangement. In the Teachers' Pets task, six students (6A, 7A, 7B, 8A, 8B, 8C) grouped the data into categories not given in the data set. For instance, instead of grouping all Dobermans or all Persian cats together, these students formed new groups: Dogs, Cats, Birds, and Reptiles (or Lizards). They went on to sum each new animal group and represent their data in a bar graph, line plot or circle graph.

Table 2 shows how students arranged the data by task, including initial and alternate arrangements. The use of an arrangement was task specific. For Shoe Size, the majority of students' ways of handling the data was split between Sequential and Summative Arrangements. With Teachers' Pets a majority of the students used either the Regrouped-Summative or Summative Arrangement. For the Academy Awards and Classical Music tasks, a majority of the students used No Arrangement as their choice of handling the data. We attribute this to a number of reasons. Both tasks included two data sets presented in an organized table. Some students were satisfied with the presentation and felt no need to rearrange the data while several students expressed their inability to rearrange two data sets within a single task. Still other students found the averages of the data sets, which did not involve reorganizing the original data sets.

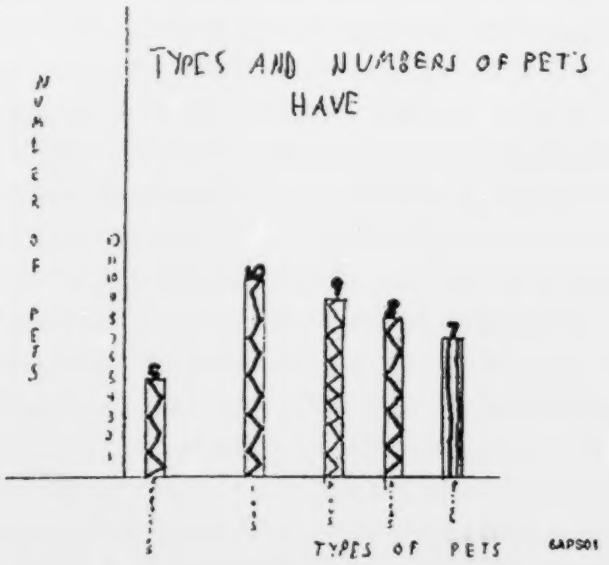

Figure 5. Regrouped-summative arrangement.

BEST COPY AVAILABLE

Table 2. Student Arrangements by Task

	Shoe Size (Task 1)	Teachers' Pets (Task 2)	Academy Awards (Task 3)	Classical Music (Task 4)
No Arrangement	4 27%	0 0%	7 41%	11 73%
Clustered Arrangement	0 0%	4 25%	3 18%	0 0%
Sequential Arrangement	6 40%	1 6%	4 24%	4 27%
Summative Arrangement	5 33%	5 31%	2 12%	0 0%
Regrouped-Summative Arrangement	0 0%	6 38%	1 5%	0 0%

Note. The percentage values represent the number of students using each arrangement per task. There were a total of 15 Shoe Size, 16 Teacher Pets, 17 Academy Awards, and 15 Classical Music arrangements.

When viewing the results of student arrangements by grade levels (see Table 3), we noticed that the highest percentages of students' arrangements found across all grades levels was No Arrangement. Recall, arrangements in this category include inappropriate arrangements as well as no attempt to handle the data. When the students selected an appropriate arrangement, both the sixth and seventh graders' highest percentage of arrangements was Sequential. The eighth grade students' highest percentage was Summative Arrangement. This grades' lowest percentage was Clustered Arrangement, which was the same for the seventh graders. The lowest arrangement percentage for the sixth graders was Regrouped-Summative.

Table 2. Student Arrangements by Grade Level

	6th Grade	7th Grade	8th Grade
No Arrangement	7 35%	5 26%	10 42%
Clustered Arrangement	3 15%	2 11%	2 8%
Sequential Arrangement	7 35%	5 26%	3 13%
Summative Arrangement	2 10%	4 21%	6 24%
Regrouped-Summative Arrangement	1 5%	3 16%	3 13%

Note. The percentage values represent the number of students using each arrangement out of a total of 20 sixth grade, 19 seventh grade, and 24 eighth grade arrangements.

Discussion

The results of this study show that a majority of the middle school students interviewed were able to reorganize, rename, and sum single sets of unorganized data and give alternate arrangements of data. However, when organized data sets were presented to the students, several were unable to interpret or rearrange the data, possibly due to the fact that they had to handle more than one data set within a single task. Many students manipulated the numbers in the data set(s) and found the average value of the data as opposed to organizing or reorganizing the data. In this study, sixth and seventh graders were more likely than eighth graders to arrange data in an appropriate manner. The eighth grade students mostly tried to find averages for every task or they felt the organized data was in the appropriate format for the school newspaper. When probed to provide a graphical representation, these students were more likely to draw a box-and-whiskers graph.

Conclusion and Recommendations

Though this research was conducted using a small sample size, it does show that students can handle statistical data when given the opportunity. The sixth grade students we interviewed had recently completed the unit on statistics and probability which would explain the percentage of students who were able to arrange the data appropriately. Retention and additional practice would explain the increase in the percentage for the seventh grade students. However, since eighth grade textbooks generally extend the curriculum to include box-and-whisker, stem and leaf graphs, and line plots, they have an extended repertoire of graphical representations. This additional information with the lack of practice could explain why the eighth grade students had a high percentage of inappropriate graphs. We conclude that students need to be presented with raw, unorganized data sets in addition to organized data sets and asked to arrange or rearrange data in a meaningful way. This practice will allow students' statistical reasoning to evolve and develop which should lead them to effectively organize, analyze, and present data as suggested by NCTM (2000).

References

Charles, R. I., Dossey, J. A., Leinwand, S. J., Seeley, C. L., & Vonder Embse, C. B. (1998). *Middle school math*. Menlo Park, CA: Scott Foresman Addison-Wesley.

Curcio, R. R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education, 18*, 382-393.

Jones, P. S., & Coxford, A. F. (Eds.). (1970). *A history of mathematics education in the United States and Canada*, Thirty-second yearbook of the National Council of Teachers of Mathematics. Washington, DC: National Council of Teachers of Mathematics.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data*. Thousand Oaks, CA: Sage.

Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal For Research in Mathematics Education, 26,* 20-39.

Mooney, E. S., Langrall, C. W., Hofbauer, P. S., & Johnson, Y. A. (2001). Refining a framework on middle school students' statistical thinking. In R. Speiser, C. A. Maher, & C. N. Walter (Eds.), *Proceedings of the twenty-third annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 437-447). Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

Paulos, J. A. (1988). *Innumeracy: Mathematical illiteracy and its consequences.* New York: Hill and Wang.

Pieters, R. S., & Kinsella, J. J. (1959). *Statistics. The growth of mathematical ideas, Grades K-12,* Twenty-fourth Yearbook of the National Council of Teachers of Mathematics (pp. 272-326). Reston, VA: National Council of Teachers of Mathematics.

Shulte, A. P., & Smart, J. R. (1981). *Teaching statistics and probability,* 1981 yearbook of the National Council of Teachers of Mathematics. Reston, VA: National Council of Teachers of Mathematics.

Steen, L. A. (Ed.). (1997). *Why numbers count: Quantitative literacy for tomorrow's America.* NewYork: College Entrance Examination Board.

Strauss, S., & Bichler, E. (1988). The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education, 19,* 64-80.

Steen, L. A. (Ed.). (1997). *Why numbers count: Quantitative literacy for tomorrow's America.* New York: College Entrance Examination Board.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

[X] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

[ ] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").